CREATING COST GROWTH MODELS FOR THE ENGINEERING AND MANUFACTURING DEVELOPMENT PHASE OF ACQUISITION USING LOGISTIC AND MULTIPLE REGRESSION

THESIS

Brandon M. Lucas, Captain, USAF

AFIT/GCA/ENC/04-02

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GCA/ENC/04-02

CREATING COST GROWTH MODELS FOR THE ENGINEERING AND MANUFACTURING DEVELOPMENT PHASE OF ACQUISITION USING LOGISTIC AND MULTIPLE REGRESSION

THESIS

Presented to the Faculty

Department of Mathematics and Statistics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Cost Analysis

Brandon M. Lucas, B.A.

Captain, USAF

March 2004

AFIT/GCA/ENC/04-02

CREATING COST GROWTH MODELS FOR THE ENGINEERING AND MANUFACTURING DEVELOPMENT PHASE OF ACQUISITION USING LOGISTIC AND MULTIPLE REGRESSION

Brandon M. Lucas, B.A.
Captain, USAF

Approved:

/Signed/
Dr. Edward D. White (Chairman) date

/Signed/
Michael A. Greiner, Major, USAF (Member) date

/Signed/
Michael J. Seibel (Member) date

AFIT/GCA/ENC/04-02

Abstract

Cost growth is a concern for all parties involved in the DoD acquisition process. These parties include cost analysts, program managers, senior DoD decision-makers, Congress, and even the American public. All of these people have a vested interest in the cost of DoD programs and most would like to see those costs decrease; as such, Congress has enacted multiple laws and reforms over the past three decades in an attempt to curb cost growth within DoD acquisition.

Previous research creates the foundation for the use of a two-step methodology to help predict cost growth, which we follow closely. First, utilizing logistic regression we analyze whether specific program characteristics predict cost growth within the Engineering and Manufacturing Development (EMD) phase for combined RDT&E and procurement budgets. The second step uses this answer (i.e., a positive response) to find cost growth predictor variables. Specifically, we perform a multiple regression analysis and determine the amount of cost growth incurred by these DoD programs. Through these two steps, we seek to unearth any predictive relationships within the data in order to build a predictive cost growth model. The final model predicts both whether a program will have cost growth and what the potential amount of the cost growth will be for the combined RDT&E and procurement budgets within the EMD phase of acquisition.

Acknowledgements

I would like to thank Dr. Edward White, my thesis advisor, for all his patience with my repetitive questions and his guidance on this project. I extend a sincere thanks to my committee members, Mr. Michael Seibel and Major Michael Greiner for their constructive guidance on this thesis. I truly appreciate the efforts of Captain John Bielecki for his continued assistance after he had already moved on to greener pastures. Finally, I would like to thank my family and friends who encourage me to never stop…even once I reach my goals.

Brandon M. Lucas

Table of Contents

Page

List of Figures

List of Tables

Table Page

CREATING COST GROWTH MODELS FOR THE ENGINEERING AND MANUFACTURING DEVELOPMENT PHASE OF ACQUISITION USING LOGISTIC AND MULTIPLE REGRESSION

I. Introduction

General Issue

Cost growth is a concern for all parties involved in the DoD acquisition process. These parties include cost analysts, program managers, senior DoD decision-makers, Congress, and even the American public. All of these people have a vested interest in the cost of DoD programs and most would like to see those costs decrease; as such, Congress has enacted multiple laws and reforms over the past three decades in an attempt to curb cost growth within DoD acquisition. Most of these efforts meet with little success and weapon systems continue to experience an average of approximately 20 percent cost growth (Drezner, 1993: xiii-xiv).

Over the past 40 years, increases in federal entitlements, such as social security and welfare, have reduced the percentage of defense outlays from 52.2 percent to 16.3 percent of total federal spending (Schick, 2000: 18). Yet, even as the funding pool evaporates, both Congress and the public demand more oversight of DoD programs. Beginning in 1969 with the Packard Initiatives, the government created a launch pad for several revisions of the acquisition regulations ("the 5000 series") and the eventual establishment of growth thresholds through the Nunn-McCurdy Act in 1982 (Christensen, 1999: 253). These thresholds serve as indicators for Congress to provide

additional oversight or to even certify that a program is necessary for national security when cost growth exceeds 15 and 25 percent respectively (Weinberger, 2002). Even with these thresholds, cost growth remains a subject of continued regulation revisions and oversight commissions.

While there are many possible causes of cost growth, and almost each one is addressed by an act or commission of some sort, a cause of certain importance is program risk. By identifying possible areas of risk within a program, program managers and cost analysts can assign dollar amounts to those risks and produce better cost estimates. Since cost growth is measured as the growth a program experiences from its initial baseline estimate to the program's current estimate, the building of a realistic and more accurate initial estimate should reduce the amount of cost growth a program encounters (Hough, 1992:v). By assigning the proper dollar amounts to a program's identified risks, cost estimators can produce more accurate cost estimates and help program managers and other senior leaders avoid cost growth and additional oversight (Sipple, 2002: 2).

Specific Issue

Multiple methods exist to conduct a cost analysis. The current stage of program development drives the choice of which method a cost estimator should use. The five most common cost estimating techniques include analogy, expert opinion, engineering, actual costs, and parametric.

Analysts primarily use analogy and expert opinion at the beginning of a program's life cycle. As many details about the project will still be unknown, estimators often attempt to compare (analogy) aspects of the current program with similar aspects from past programs whose costs are known. Similarly, subject matter experts will

frequently be called upon to advise the cost estimator about certain portions of the project and provide their own estimates for potential cost. Naturally, such subjective techniques may be relatively imprecise, but they can provide fast, inexpensive and easily modifiable estimates during initial periods of the program.

Estimators generally use the engineering and actual costs methods in the later stages of program development in order to produce more detailed cost estimates. The engineering technique actually constructs the cost estimate beginning with the lowest element within the work breakdown schedule and works its way to the top. As such, it consumes a lot of time and costs more than other techniques, leaving little opportunity for "what-if" drills, but produces quite objective and accurate estimates. Likewise, the actual costs method extrapolates the data gathered over the course of the program to produce a cost estimate. While this technique creates very objective and accurate reports, the findings may be of limited use because the program is in its later stages.

The fifth method, parametric, is also known as the statistical method. When estimators use this technique, they employ one or more databases of comparable programs to formulate statistical inferences about the relationships between the programs. While it is important to note that the databases' range of parameters limit the applicability of the findings, the findings do represent a relatively objective, inexpensive and modifiable method for creating cost estimates. In fact, some cost estimators believe these advantages represent a superior technique to the other methods listed above and through their collective efforts these researchers provide a basis for further study.

Presented with these facts, we use parametric techniques to construct cost growth models. Specifically, this work builds upon that of Sipple (2002), Bielecki (2003) and

Moore (2003) to create a predictive model for cost estimators. By using statistical regression, both cost growth relationships and the amount of cost growth can be predicted (Sipple, 2002:2). Moreover, by utilizing such a model early in the program's life cycle, more accurate cost estimates can be produced and cost growth can be decreased.

Scope and Limitations of the Study

To generate such a predictive model, an analyst needs access to a proper database. The Selected Acquisition Reports (SARs) represent an invaluable source of information to populate such a database. Congress mandates the production of these annual reports by the individual weapon system program offices and also sets the formatting standards and funding thresholds— two key aspects that make the reports more useable for our purpose. Part of this standardization requires certain information, which provides data for many of the possible predictor variables, to be reported as well.

Specifically, the SARs report values for the planning estimate (PE), development estimate (DE) and production estimate (PdE) (if available). Also, the reports provide the current estimate (CE), which serves as the most recent estimate for the program. The cost data further divides into sections for Research and Development, Test and Evaluation (RDT&E), procurement and military construction (Jarvaise, 1996:3). While differing organizations use these three estimates in various ways to calculate cost variance, the method we use in this study calculates the difference between the CE and DE. Therefore, in accordance with aforementioned preceding works, "we define cost variance as the difference between the Current Estimate to the Development Estimate and cost growth as positive cost variance" (Moore, 2003:4).

The SARs further separate cost variance into seven categories based on program effects: Estimating (or Escalation), Quantity, Schedule, Engineering, Support, Economic and Other (Jarvaise, 1996:4). These categories report both base-year and then-year dollars for the program and thus allow researchers to account for inflation. Through these unique divisions, we compare cost variance throughout the database and search for possible predictor variables of cost growth.

For this research, we only analyze programs that use the DE as the baseline estimate and only include the current SAR for the program. This study compiles the efforts of Sipple (2002), Bielecki (2003) and Moore (2003) by examining cost growth in both RDT&E and procurement within the Engineering and Manufacturing Development (EMD) phase of acquisition. Like the preceding authors, we include only five of the seven cost variance categories and exclude both Economic and Quantity cost from the analysis as these categories are outside of the estimator's realm of control (Bielecki, 2003:4). Therefore, this study only examines cost variance for Estimating, Engineering, Schedule, Support, and Other for both RDT&E and procurement funding within the EMD phase.

As previously mentioned, the database itself limits this research. By using the SAR data, limits already exist due to security classification and unknown budgeting for risk (Sipple, 2002:4). Thus, we cannot include some data in this study due to its security classification, while some estimates (i.e., DE) may include unknown dollar amounts by program managers in an attempt to budget, or hedge, for risk. We address these issues in full in Chapter III.

While this research differs from most prior DoD research by being inferential rather than descriptive, the precedent for its use and applicability has already been set by Sipple (2002), Bielecki (2003), and Moore (2003). This research again utilizes the two-step regression methodology Sipple developed; first, logistic regression analysis predicts which programs will have cost growth, and then multiple regression analysis predicts how much cost growth the program will incur. Sipple (2002) provides the initial groundwork by testing the two-step process with only Engineering cost data within RDT&E. Bielecki (2003) further validates the process by using the remaining cost categories within RDT&E. Finally, Moore (2003) uses the same process, but performs his analysis on the procurement funding within the EMD phase. This research only differs from the other three studies in that it combines each of these three areas and produces an overall model for the EMD acquisition phase.

Research Objectives

This study has two main objectives. First, utilizing logistic regression we analyze whether specific program characteristics predict cost growth within the EMD phase for combined RDT&E and procurement budgets. "Logistic regression differs from multiple regression in that it predicts a binary response. In our case, the binary response is: *Does a program experience cost growth, Yes or No* (Sipple, 2002:5)?" The second objective uses this answer (i.e., a positive response) to find cost growth predictor variables. Specifically, we run a multiple regression analysis and determine the amount of cost growth incurred by the significant program characteristics. Through these two objectives, we seek to unearth any predictive relationships within the data in order to build a predictive model. The final model predicts both whether a program will have cost

growth and what the potential range of the cost growth will be for the combined RDT&E and procurement budgets within the EMD phase of acquisition.

Chapter Summary

This research builds directly off of the contributions made by Sipple (2002), Bielecki (2003) and Moore (2003). The end result of this work is a predictive model cost estimators can use to help account for risk and reduce cost growth within a program. By constructing a database from the SARs, we are able to utilize both logistic and multiple regression to build a model that both identifies programs that may encounter cost growth and predicts the amount of cost growth. The result is a tool program managers and cost estimators can use to identify problems early on within a program, which may help control and reduce the amount of measured cost growth. To do so, in the subsequent chapters we provide a review of pertinent cost estimating literature, a detailed synopsis of our methodology, an analysis of our findings and results, and our conclusions from this research.

II. Literature Review

Chapter Overview

This chapter reviews key factors and events that contribute to the field of cost estimating. Specifically, the review discusses recent developments in the acquisition environment, assesses cost estimating and cost risk, and summarizes past research related to cost growth. However, as Sipple (2002), Bielecki (2003) and Moore (2003) thoroughly reviewed this topic, we limit the scope of this chapter to a short review of key acquisition system points and a discussion of any pertinent research-related findings since January of 2003. The information from these studies provides a basis for understanding cost factors and helps to build a regression model that may predict cost growth for the Engineering and Manufacturing Development (EMD) phase of acquisition.

The Current Acquisition Setting

To best appreciate the complexities involved in estimating cost growth within DoD, the current acquisition environment must be understood. Through this governing setting, we determine both where and how cost growth occurs, as well as how the DoD measures cost growth. As with many governmental processes, the world of acquisition changes constantly, but the source of guidance remains constant: the Department of Defense Instruction (DoDI) 5000.2, Operation of Defense Acquisition System. Though significantly revised over the past two years, this instruction still shapes acquisition structure, policy, and the processes for making war fighting requirements a reality.

Specifically, at the beginning of 2001, revisions to the DoDI 5000.2 reduced the required milestones from four to three; the four previous milestones known as MS 0, MS

I, MS II, and MS III are now labeled as A, B, and C (DoDI 5000.2). These milestones, or major dividers within a program's development, serve as decision points within the acquisition process, both for program review and fiscal purposes. The milestone change uniquely impacts the latter purpose because the milestones determine the phase of the program, as well as impact the cost estimates. The previous four phases follow: Phase 0 – Concept Exploration, Phase I – Program Definition and Risk Reduction, Phase II – Engineering and Manufacturing Development, and Phase III – Production, Fielding/Deployment, and Operation Support (DoDI 5000.2). The three classifications of activities are Pre-System Acquisition, System Acquisition, and Sustainment. For further clarification of this taxonomy, a thesis entitled *Correlation Analysis: Army Acquisition Program Cycle Time and Cost Variation* by Howard Jaynes (1999: 11-13; Bielecki, 2003: 8-9) provides the following concise summary of each milestone and phase.

- Milestone 0: conduct concept studies. Validation of the mission need and identification of possible alternatives. Approval of MS 0 by the Defense Acquisition Board authorizes entry into Phase 0.
- Phase 0: Concept Exploration. The mission need and the alternatives are further defined in terms of cost, schedule, and performance objects. Costs are incorporated in the Acquisition Program Baseline (APB). Acquisition Strategies are developed and the Operation Requirements Document is prepared.
- Milestone I: official approval to begin a new program.
- Phase I: Program Definition and Risk Reduction. The program is defined in terms of designs and technological approaches. Prototyping and early operational assessments are used to reduce risk. Identification of cost and schedule trade-offs.

- Milestone II: approval to enter Phase II. The Milestone Decision Authority (MDA) evaluates the acquisition strategy and updated APB (development baseline) of the program before authorizing continuation. Note: this is the estimate we use in our research to calculate cost growth.
- Phase II: Engineering and Manufacturing Development. The program is transformed into a cost-effective, stable design. Developmental testing is conducted to ensure performance capabilities are satisfied and Low Rate Initial Production is authorized to further validate the new system.
- Milestone III: approval to enter Phase III. MDA reviews the acquisition strategy and updated APB (production baseline) program before approving entry into Phase III.
- Phase III: Production, Fielding/Deployment and Operational Support. The program enters full rate production and works to achieve Initial Operational Capability (IOC). IOC is the first deployment of a weapons system to an operational unit.

However, as mentioned earlier, beginning in 2001 the new classification took effect. Unlike the previous studies by Sipple (2002), Bielecki (2003), and Moore (2003), our data includes the period from 1990 – 2002, or one more fiscal year, which indicates that our Selected Acquisition Report (SAR) data may be more affected by this change. However, we find that relatively few of the programs in the SAR are affected by the new milestone strategy and decide to focus only on programs that use the previous methodology. As such, our data remains more consistent and less affected by the potential changes the new milestone strategy could initiate.

Moreover, to help conceptualize these differing milestones, phases, and where our research lies within them, we also provide the chart, Figure 1, below. As our definition of cost growth is based on the percentage price increase from the Development Estimate (DE) to the Current Estimate, we find it beneficial to illustrate where our small portion of research fits within the acquisition framework.

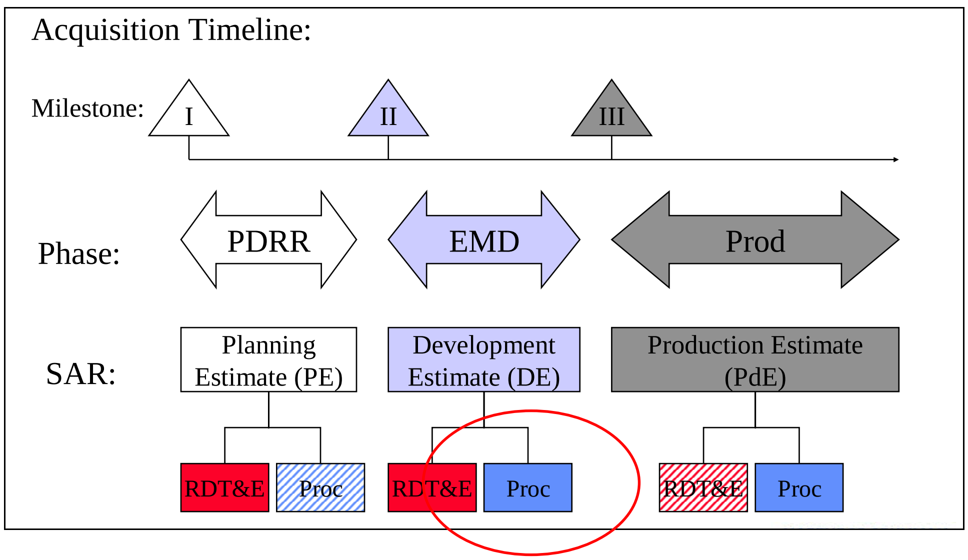


Figure 1 – Acquisition Timeline (Dameron, 2001: 4)

The Cost Estimating Process and Risk Assessment

As all three concepts are intertwined, any discussion concerning cost growth also requires knowledge about cost estimation and risk assessment. In the first chapter of this study we define cost growth as "the growth a program experiences from its initial baseline estimate to the program's current estimate," which means that cost growth equals the actual amount of funds the program goes over budget. Similarly, risk assessment establishes a monetary amount for cost risk, with cost risk being the predicted

dollar amount of cost growth likely to incur in a program (Coleman, 2000:3). Since cost growth closely relates to cost risk, a review of risk assessment methods helps us to understand how experts measure the price of possible cost growth. Figure 2 shows a chart of risk methods the analysts within the Ballistic Missile Defense Organization (BMDO) commonly utilize for their estimates (Coleman, 2000:4; Sipple, 2002:17).

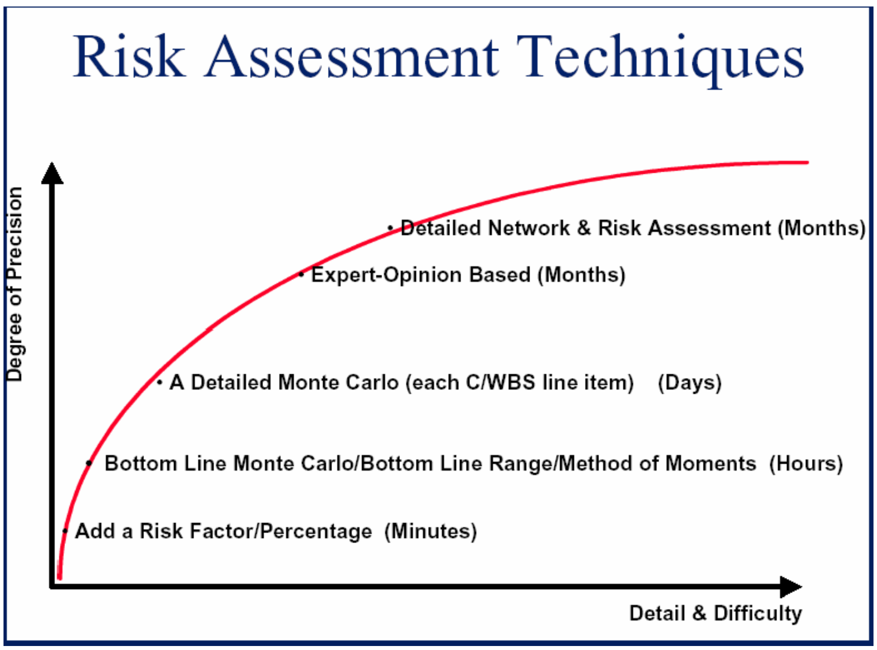


Figure 2 – Risk Assessment Techniques (Coleman, 2000: 4-9)

While not necessarily exhaustive, the chart does cover many of the major techniques risk estimators use and also serves as a good foundation for general discussion about measuring risk. A succinct definition of each technique is provided by Bielecki (2003: 14-15), based on a cost analysis symposium (Coleman, 2000: 4-9, 12, 16), below:

- Add a Risk Factor/Percentage: is the least precise and easiest technique to use. Relies on technical expert judgment to assign a high-level, subjective risk factor for the estimate.
- Bottom Line Monte Carlo / Bottom Line Range / Method of Moments: may use Monte Carlo Simulation, but on higher levels of the Work Breakdown Structure (WBS). Other uses include a limited database, analogy methodology or expert opinion to determine risk estimates.
- Detailed Monte Carlo Simulation: C/WBS is the Cost or Work Breakdown Structure. Uses Monte Carlo Simulation, but relies on historical data to develop probability distributions of cost outcomes.
- Expert-Opinion Based: relies on surveys of experts to determine the possible distributions of WBS item costs. Uses Monte Carlo simulation to estimate a range of possible costs. Assumes experts are accurate.
- Detailed Network & Risk Assessment: is the most precise and most difficult to apply. It requires a very detailed schedule and task breakout. It uses a beta or triangular distribution to schedule item durations and creates a stochastic model from which to estimate the risk of a schedule slip. The estimator uses the Monte Carlo Simulation method to estimate the cost.

Past Research in Cost Growth

The last portion of this literature review examines past research dealing with cost growth. Due to congressional and DoD emphasis on accurate cost estimates and within-budget acquisition programs, much research exists pertaining to cost growth. However, as this research directly follows to the research of Sipple (2002), Bielecki (2003) and

Moore (2003), only a cursory review is provided; Sipple laid an exhaustive and solid foundation upon which both Bielecki and Moore built, to which we only either summarize or update with any recent (i.e., 2003-2004) findings.

This prior research assists us immensely in our cause as it provides credible explanations for the basic sources of cost growth. Indeed, Sipple first establishes the copious list of predictor variables that later researchers use to build predictive regression models. Table 1 below lists the many studies Sipple reviewed.

Table 1 – Sipple Thesis (Sipple, 2002:20-44)

Author (Year)
IDA (1974)
Woodward (1983)
Obringer (1988)
Singleton (1991)
Wilson (1992)
RAND (1993)
Terry & Vanderburgh (1993)
BMDO (2000)
Christensen & Templin (2000)
Eskew (2000)
NAVAIR (2001)
RAND (2001)

This list does not exhaust the studies previously examined. In another thesis, Gordon (1996) provides a listing of studies on cost growth performed both by the RAND Corporation and the Air Force Institute of Technology. Bielecki (2003) provides two tables summarizing both of these lists by Gordon. See Tables 2 and 3 below.

Table 2 – RAND Reports (Gordon, 1996:2-2)

Author (Year)	Findings	Sensitivity Factors
Jarvaise, et al. (1996)	Defense System Cost Performance Database	Derived from SARs
Drezner, et al. (1993)	Cost Estimates biased toward underestimation by about 20% from PE and DE and 2% from PdE	Program Size, Maturity
Drezner (1992)	No demonstrated relationship between prototyping and cost or schedule outcomes (67)	No Program Phase, Not System Type
Hough (1992)	Selected Acquisition Reports can Delay, Mask or Exclude Significant Cost Growth	Economic, Quantity, Schedule, Engineering, Estimating and Other Changes

While our review remains concise and only references the aforementioned works, one study does need to be mentioned in more detail. The RAND study from 1993 utilizes SAR data for its tests. Based on the conclusions of the study, RAND finds that inflation and quantity, two of the seven cost variances listed in Chapter I, have the largest impact on cost growth. However, due to the nature of a cost estimate, which already includes and accounts for these factors, RAND establishes that the two factors can be excluded when analyzing for cost growth. Thus, Sipple (2002) adopts this principle of exclusion and our research methodology also excludes inflation and quantity.

Table 3 – AFIT Theses (Gordon, 1996:2-3)

Author (Year)	Findings	Sensitivity Factors
Nystrom (1996)	Complex non-linear EAC methods not superior to simpler index based EAC methods	Stage of Completion, System Type, Program Phase, Contract Type, Service Component, and Inflation
Buchfeller and Kehl (1994)	No Significant Differences in Cost Variances between categories	Not Service, Not Program Phase, Not Contract Type, Not Stage of Completion
Elkinton and Gondeck (1994)	BAC Adjustment Factors derived from Historical "Cost Growth" do not Improve EACs	Not Contract Type, Not Stage of Completion
Pletcher and Young (1994)	Contracts which Improved Cost Performance over time differ from those which Worsen	Performance Management Baseline Stability
Terry and Vanderburgh (1993)	SCI based EAC best predictor of CAC for all Stages of Contract Completion	Contract Completion Stage, Program Phase, Contract Type, Service Component, System Type, Major Baseline Changes, but not Management Reserve
Wandland (1993)	Completed Contracts have more "Cost Growth" than Sole Source	Not Contract Type, Not Absolute Price
Wilson (1992)	Cost Overruns at Completion are Worse than between 15 and 85% complete (=.15)	Service (except Navy), Contract Type, System Type, and Program Phase, but not relative time
Singleton (1991)	"Cost Growth" can be predicted based on three factors	Schedule Risk, Technical Risk and Configuration Stability
Obringer (1988)	"Cost Growth" is not attributable to increased Industry Direct or Overhead to Total Cost Ratio	Specific Contractors (8 of 16) showed growth between 1980 and 1986
Blacken (1986)	"Cost Growth" varies with Characteristics of Contract Changes	Scope, Number of Effected SOW Pages, Contract Type, Change Type, Time to Definitize, Time to Negotiate, Not to Exceed Estimate, Stage of Completion, Stage of Development, Schedule Changes, Length of ECP, Length of Period of Performance

Finally, we would be remiss not to detail some of the findings from the forerunners of this research: Sipple (2002), Bielecki (2003) and Moore (2003). Sipple's research provides the basis for this research through the creation of a new framework and adopted methodology, as well as by creating the list of predictor variables (Chapter III contains the complete list). Moreover, Sipple's use of both logistic and multiple regression for predictive model building acts as a forerunner for the two-step methodology based on the findings of our research within the cost arena. Narrow in focus due to the study's groundbreaking approach, Sipple (2002) analyzes only RDT&E funding for the engineering category within the EMD phase of acquisition. Using the same methodology, Bielecki (2003) studies the remaining four categories of RDT&E within EMD. Moore (2003) again looks at EMD, but rather than RDT&E funding he examines procurement funding only. As such, these three studies lay the foundation for our current research, the analysis of both RDT&E and procurement funding within the EMD stage of acquisition.

Chapter Summary

In this chapter, we define how the current acquisition environment and cost estimating process relate to and affect this research. Also, this chapter provides a brief outline of past research in cost growth by acknowledging the extensive literature review performed by Sipple (2002) that serves as the basis for future cost growth research. Using the predictor variables past studies identify, we continue the process of building a predictive model for cost growth within the EMD phase. In so doing, we build the foundation for our research methodology set forth in the next chapter.

III. Methodology

Chapter Overview

This chapter clarifies the methodology by which we conduct this research. We begin by examining the database we use in more detail to elucidate both its advantages and disadvantages. We continue with a brief overview of two key procedural aspects: data collection and candidate variable compilation. We conclude the chapter by reviewing exploratory data analysis and detailing the regression techniques we use.

Database Characteristics

As mentioned in previous chapters, we use the Selected Acquisition Reports (SAR) to build our database for this study. Each SAR contains a diverse array of data to include a narrative program summary, schedule and budget information, cost variances and performance to name a few. The reports present this data in both base year and then year dollars, with base year dollars as our preferred choice. We use base year (BY) dollars and adjust these amounts to a standard base year (BY 2002) so meaningful comparisons can be made amongst the data. Furthermore, each report breaks the cost variance data into the seven different cost categories:

- Economic: changes in price levels due to the state of the national
- Quantity: changes in the number of units procured
- Estimating: changes due to refinement of estimates
- Engineering: changes due to physical alteration
- Schedule: changes due to program slip/acceleration
- Support: changes associated with support equipment

- Other: changes due to unforeseen events (Drezner, 1993:7)

This research uses the sum of these categories (excluding economic and quantity) for both the RDT&E and procurement funding amounts annotated in the cost variance summaries. As such, it represents a logical growth from the three previous studies by Sipple (2002), Bielecki (2003), and Moore (2003), which looked at separate funding components within the EMD phase of acquisition.

Specific Limitations of SAR Data

By using the SARs for our database, we immediately encounter one limitation; Congress only requires SARs for ACAT IC and D programs (Knoche, 2001:1), and therefore we construct our predictive models with these specific programs. Designated by dollar thresholds, ACAT IC and D represent large programs, which limit our study by disregarding the multitude of smaller DoD programs. While our database captures many key military programs, the database does not account for most programs.

Other problems occasionally occur due to the use of SAR data as well. While laying the foundation for this research, Sipple (2002) discovers that SAR data contains some possible disadvantages. While the list below by Hough (1992) only represents a summary of these potential difficulties, both Sipple (2002) and Bielecki (2003) provide a more exhaustive review of these problems and what effects they can wreak on a database.

- Failure of some programs to use a consistent baseline cost estimate
- Exclusion of some significant elements of cost
- Exclusion of certain classes of major programs (e.g., special access programs)
- Constantly changing preparation guidelines
- Inconsistent interpretation of preparation guidelines across programs

- Unknown and variable funding levels for program risk
- Cost sharing in joint programs
- Reporting of effects of cost changes rather than their root causes (Hough, 1992:v)

Data Collection

We inherit our database from Sipple (2002), Bielecki (2003), and Moore (2003). Sipple's original database includes fields for both RDT&E and procurement funding, and consists of all available SARs from 1990 to 2000 that use the development estimate (DE) as the baseline estimate. Bielecki and Moore update the database in 2003 to include data through December 2001. Our database benefits from the passage of time even more so and adds an additional year—2002.

However, we do not only add information for 2002, but additionally complete a thorough review of the entire database by comparing the latest SAR using an EMD-based DE with the information already in the database. By "scrubbing" the database, we reduce the possibility of human input error and also attain a better understanding of the complex task at hand. Finally, like our predecessors, we include joint service programs and exclude program information that has a security classification other than "unclassified."

Exploratory Data Analysis

Before building any models, the data must be understood so that we employ the proper procedures. Our tests reveal the same mixed distribution Sipple (2002) encounters when he plots the response variable. Specifically, two types of data comprise our response data: continuous and discrete. As in previous cases, the discrete data centers at zero, while the continuous data is sporadically spaced throughout the continuum. Since probability theory requires that the chance of obtaining a specific value within a

continuous distribution is zero, and the discrete mass centered at zero nullifies that probability, the normal solution entails splitting the data into two sets. To accomplish this task however, we use the two-step methodology by Sipple (2002). We find that Moore (2003:22) provides an excellent summation of Sipple's two-step method:

> We first split the data into discrete and continuous distributions. We then utilize logistic regression to analyze the discrete distribution and multiple regression to analyze the continuous distribution. Thus, we develop two models: a logistic regression model that analyzes the full data set to predict whether or not a program will have procurement cost growth, and a multiple regression model that analyzes only programs containing procurement cost growth to predict the amount of cost growth we expect. For the logistic regression portion of our analysis, we convert all negative cost growth to zero cost growth. Furthermore, to ensure that we construct a robust model, we set approximately 20 percent of our data aside for validation before we begin any regression analysis (Sipple, 2002: 59).

We then randomly select 20 percent of the data and set it aside until needed for model validation. However, before actually doing logistic or multiple regression we must create both the response and candidate predictor variables.

Response Variables

Our research aims to build both a logistic and a multiple regression model to predict cost growth for the EMD phase of acquisition, which consists of both RDT&E and procurement funding. While previous studies have not combined the cost variance categories of RDT&E and procurement, our research follows the same basic approach and likewise requires two different response variables. The first variable, *EMD Cost Growth?*, indicates whether or not cost growth occurs within a given program. As the response only answers yes or no, we choose to use a binary variable where '1' indicates that a program experiences cost growth and '0' indicates that the program does incur cost growth.

The other variable, *EMD %*, takes the form of a percentage rather than the actual dollar amount so that the cost growth remains relative to the program. Again, using a percentage readily allows for comparison amongst the programs, removing our need to modify the results before an equal comparison can be made. This response variable represents the percentage of cost growth within the EMD phase by both RDT&E and procurement funding.

Predictor Variables

Our research utilizes the same basic pool of candidate variables established by Sipple (2002). Through Sipple's in-depth research, these contenders prove themselves to be capable predictors. As such, they provide a solid pool of variables from which to select for model building.

These variables readily divide into five categories: program size, physical type of program, management characteristics, schedule characteristics, and other characteristics. As some of these categories are general in nature, Sipple also develops sub-categories where needed. The list below categorizes these candidates and provides a concise description for better comprehension (Sipple, 2002:61).

Program Size Variables

- *Total Cost CY $M 2003* – continuous variable which indicates the total cost of the program in CY $M 2003
- *Total Quantity* – continuous variable which indicates the total quantity of the program at the time of the SAR date; if no quantity is specified, we assume a quantity of one (or another appropriate number) unless the program was terminated
- *Prog Acq Unit Cost* – continuous variable that equals the quotient of the total cost and total quantity variables above
- *Qty during PE* – continuous variable that indicates the quantity that was estimated in the Planning Estimate
- *Qty planned for R&D$* – continuous variable which indicates the quantity in the baseline estimate

Physical Type of Program

- Domain of Operation Variables
  - *Air* – binary variable: 1 for yes and 0 for no; includes programs that primarily operate in the air; includes air-launched tactical missiles and strategic ground-launched or ship-launched missiles
  - *Land* – binary variable: 1 for yes and 0 for no; includes tactical ground-launched missiles; does not include strategic ground-launched missiles
  - *Space* – binary variable: 1 for yes and 0 for no; includes satellite programs and launch vehicle programs
  - *Sea* – binary variable: 1 for yes and 0 for no; includes ships and ship-borne systems other than aircraft and strategic missiles
- Function Variables
  - *Electronic* – binary variable: 1 for yes and 0 for no; includes all computer programs, communication programs, electronic warfare programs that do not fit into the other categories
  - *Helo* – binary variable: 1 for yes and 0 for no; helicopters; includes V-22 Osprey
  - *Missile* – binary variable: 1 for yes and 0 for no; includes all missiles
  - *Aircraft* – binary variable: 1 for yes and 0 for no; does not include helicopters
  - *Munition* – binary variable: 1 for yes and 0 for no
  - *Land Vehicle* – binary variable: 1 for yes and 0 for no
  - *Ship* – binary variable: 1 for yes and 0 for no; includes all watercraft
  - *Other* – binary variable: 1 for yes and 0 for no; any program that does not fit into one of the other function variables

Management Characteristics

- Military Service Management
  - *Svs > 1* – binary variable: 1 for yes and 0 for no; number of services involved at the date of the SAR
  - *Svs > 2* – binary variable: 1 for yes and 0 for no; number of services involved at the date of the SAR
  - *Svs > 3* – binary variable: 1 for yes and 0 for no; number of services involved at the date of the SAR
  - *Service = Navy Only* – binary variable: 1 for yes and 0 for no
  - *Service = Joint* – binary variable: 1 for yes and 0 for no
  - *Service = Army Only* – binary variable: 1 for yes and 0 for no
  - *Service = AF Only* – binary variable: 1 for yes and 0 for no
  - *Lead Svc = Army* – binary variable: 1 for yes and 0 for no
  - *Lead Svc = Navy* – binary variable: 1 for yes and 0 for no
  - *Lead Svc = DoD* – binary variable: 1 for yes and 0 for no
  - *Lead Svc = AF* – binary variable: 1 for yes and 0 for no
  - *AF Involvement* – binary variable: 1 for yes and 0 for no
  - *N Involvement* – binary variable: 1 for yes and 0 for no

- *MC Involvement* – binary variable: 1 for yes and 0 for no
  - *AR Involvement* – binary variable: 1 for yes and 0 for no
- Contractor Characteristics
  - *Lockheed-Martin* – binary variable: 1 for yes and 0 for no
  - *Northrop Grumman* – binary variable: 1 for yes and 0 for no
  - *Boeing* – binary variable: 1 for yes and 0 for no
  - *Raytheon* – binary variable: 1 for yes and 0 for no
  - *Litton* – binary variable: 1 for yes and 0 for no
  - *General Dynamics* – binary variable: 1 for yes and 0 for no
  - *No Major Defense KTR* – binary variable: 1 for yes and 0 for no; a program that does not use one of the contractors mentioned immediately above = 1
  - *More than 1 Major Defense KTR* – binary variable: 1 for yes and 0 for no; a program that includes more than one of the contractors listed above = 1
  - *Fixed-Price EMD Contract* – binary variable: 1 for yes and 0 for no

Schedule Characteristics

- RDT&E and Procurement Maturity Measures
  - *Maturity (Funding Yrs complete)* – continuous variable which indicates the total number of years completed for which the program had RDT&E or procurement funding budgeted
  - *Funding YR Total Program Length* – continuous variable which indicates the total number of years for which the program has either RDT&E funding or procurement funding budgeted
  - *Funding Yrs of R&D Completed* – continuous variable which indicates the number of years completed for which the program had RDT&E funding budgeted
  - *Funding Yrs of Prod Completed* – continuous variable which indicates the number of years completed for which the program had procurement funding budgeted
  - *Length of Prod in Funding Yrs* – continuous variable which indicates the number of years for which the program has procurement funding budgeted
  - *Length of R&D in Funding Yrs* – continuous variable which indicates the number of years for which the program has RDT&E funding budgeted
  - *R&D Funding Yr Maturity %* – continuous variable which equals *Funding Yrs of R&D Completed* divided by *Length of R&D in Funding Yrs*
  - *Proc Funding Yr Maturity %* – continuous variable which equals *Funding Yrs of R&D Completed* divided by *Length of Prod in Funding Yrs*

        - *Total Funding Yr Maturity %* – continuous variable which equals *Maturity (Funding Yrs complete)* divided by *Funding YR Total Program Length*
- EMD Maturity Measures
    - *Maturity from MS II in mos* – continuous variable calculated by subtracting the earliest MS II date indicated from the date of the SAR
    - *Actual Length of EMD (MS III-MS II in mos)* – continuous variable calculated by subtracting the earliest MS II date from the latest MS III date indicated
    - *MS III-based Maturity of EMD %* – continuous variable calculated by dividing *Maturity from MS II in mos* by *Actual Length of EMD (MS III-MS II in mos)*
    - *Actual Length of EMD using IOC-MS II in mos* – continuous variable calculated by subtracting the earliest MS II date from the IOC date
    - *IOC-based Maturity of EMD %* – continuous variable calculated by dividing *Maturity from MS II in mos* by *Actual Length of EMD using IOC-MS II in mos*
    - *Actual Length of EMD using FUE-MS II in mos* – continuous variable calculated by subtracting the earliest MS II date from the FUE date
    - *FUE-based Maturity of EMD %* – continuous variable calculated by dividing *Maturity from MS II in mos* by *Actual Length of EMD using FUE-MS II in mos*
- Concurrency Indicators
    - *MS III Complete* – binary variable: 1 for yes and 0 for no
    - *Proc Started based on Funding Yrs* – binary variable: 1 for yes and 0 for no; if procurement funding is budgeted in the year of the SAR or before, then = 1
    - *Proc Funding before MS III* – binary variable: 1 for yes and 0 for no
    - *Concurrency Measure Interval* – continuous variable which measures the amount of testing still occurring during the production phase in months; actual IOT&E completion minus MS IIIA (Jarvaise, 1996:26)
    - *New Concurrency Measure %* – continuous variable which measures the percent of testing still occurring during the production phase; (MS IIIA minus actual IOT&E completion in moths) divided by (actual minus planned IOT&E dates) (Jarvaise, 1996:26)

Other Characteristics

- *# Product Variants in this SAR* – continuous variable which indicates the number of versions included in the EMD effort that the current SAR addresses
- *Class – S* – binary variable: 1 for yes and 0 for no; security classification Secret
- *Class – C* – binary variable: 1 for yes and 0 for no; security classification Confidential
- *Class – U* – binary variable: 1 for yes and 0 for no; security classification Unclassified

- *Class at Least S* – binary variable: 1 for yes and 0 for no; security classification is Secret or higher
- *Risk Mitigation* – binary variable: 1 for yes and 0 for no; indicates whether there was a version previous to SAR or significant pre-EMD activities
- *Versions Previous to SAR* – binary variable: 1 for yes and 0 for no; indicates whether there was a significant, relevant effort prior to the DE; a pre-EMD prototype or a previous version of the system would apply
- *Modification* – binary variable: 1 for yes and 0 for no; indicates whether the program is a modification of a previous program
- *Prototype* – binary variable: 1 for yes and 0 for no; indicates whether the program had a prototyping effort
- *Dem/Val Prototype* – binary variable: 1 for yes and 0 for no; indicates whether the prototyping effort occurred in the PDRR phase
- *EMD Prototype* – binary variable: 1 for yes and 0 for no; indicates whether the prototyping effort occurred in the EMD phase
- *Did it have a PE* – binary variable: 1 for yes and 0 for no; indicates whether the program had a Planning Estimate
- *Significant pre-EMD activity immediately prior to current version* – binary variable: 1 for yes and 0 for no; indicates whether the program had activities in the schedule at least six months prior to MSII decision
- *Did it have a MS I* – binary variable: 1 for yes and 0 for no
- *Terminated* – binary variable: 1 for yes and 0 for no; indicates if the program was terminated

However, we choose to revise the list somewhat through the modification, deletion, and addition of differing variables. We make these changes in an attempt to improve both the nomenclature of certain variables as well as increase the predictive capability of our models through the new variables' inclusion. The following list documents these changes:

- Delete *Domain of Operation* – Air/Sea/Land/Space binary variables make this redundant
- Delete *Proc Cost Growth* because it includes all seven categories of cost growth; only five are needed
- Delete *Class S-R* – all of the SARs are classified secret or lower, and the variable duplicates *Class S*
- Delete *Is MSIII Complete?* – always zero since MSIII cannot be complete for our programs
- Delete *RAND Concurrency Measurement Interval and RAND Concurrency Measurement Interval %* - does not apply to programs in the EMD phase

- Delete *Terminated?* – our research applies to a living program; the variable is not applicable if the program is terminated and the variable cannot be used if the program still operates
- We delete the following group of variables for lack of data points (less than 30 would remain after we remove the 20 percent validation set):
    - *FOT&E End Planned*
    - *FOT&E End Current Estimate*
    - *MSIIIa Planned & Current Estimate*
    - *MSIIIb Planned & Current Estimate*
    - *FUE Planned*
    - *FUE Current Estimate*
    - *Maturity from MSII (current calculation in months)*
    - *Qty in PE*
- Add *LRIP Planned?* – binary with 1 for yes and 0 for no to indicate whether the program has Low Rate Initial Production
- Add *Space (RAND)* – missing from the original database, but needed for full accountability of the included programs
- Change of variable name:
    - *Qty Planned for R&D$* to *Qty Planned for R&D*
    - *Earliest Actual MSII Date* to *Current Actual MSII Date*
    - *Earliest Actual MSIII Date* to *Current Actual MSIII Date*
    - *Actual Length of EMD using (E-B)* to *Time from MSII to IOC in months*
    - *Program Acquisition Unit Cost* to *Unit Cost*
    - *Maturity of EMD using IOC* to *Maturity of EMD at IOC* (also corrected the formula so that if IOC occurs after MSIII, the percentage cannot go over 100%)

Logistic Regression

Since we first want to predict whether or not cost growth occurs, which is a 'yes' or 'no' (1 or 0) question, we choose to use logistic regression. Logistic regression by design primarily predicts a binary outcome, which suits our goal perfectly. To utilize logistic regression, we code all of the programs that have either no cost growth or negative cost growth as a '0'. We code programs with negative cost growth the same as programs with no cost growth because we are not interested in predicting negative cost growth. We then code the remaining programs, which all have positive cost growth, as a '1'. As we now have a distribution, we characterize this variable, *EMD (overall) Cost*

Growth?, as a Bernoulli random variable with probability p of success (success = 1) (Neter, 1996:568).

We base our research procedure on that of Sipple (2002), but we do make some changes. Like our predecessors, we use JMP® 4.0.4 and 5.0.1 (SAS Institute, 2001 and 2003) to compute thousands of regressions and record the results on spreadsheets. Specifically, we document the p-values, receiver operating characteristics, and R-square U values that JMP® provides. We begin with one-predictor models of all variables and select those with a p-value less than .05. We then run this selection against all of the predictor variables, and choose the top ten models as identified by having the lowest cumulative p-value. We then take the top ten and run them against all of the predictor variables. At this point, we look to decrease our variable pool by identifying any variables that do not seem to contribute to the model building process in a statistically significant way (i.e., produce a cumulative p-value model less than .1). We continue this process until the addition of another variable makes the top ten models exceed the cumulative p-value cutoff of .1. Next, we run our final models against all of the predictor variables to ensure we do not miss a statistically relevant combination. Afterwards, we try to improve our models through the use of higher order terms (e.g., squaring, cubing, natural log, inverse). Finally, we analyze the resulting candidate models to find our 'best' model, which we validate using the validation data we set aside before running the regressions.

Multiple Regression

To predict the percentage amount of cost growth, we use multiple regression to build our models. As with logistic regression, we use JMP® to compute our models. Moreover, we use the same process for building our multiple regression models as for our logistic regression models. However, we build this model using only the programs that incur cost growth (i.e., coded as a '1'), unlike the logistic model which uses all of the available programs. Finally, before we actually build the models, we apply a log transformation to our dependent, or response, variable to correct for heteroskedasticity in the residual plot (Sipple, 2002: 72). We explain the reasons for making this transformation in the next chapter.

Chapter Summary

In this chapter we explain the procedures we use to build our regression models. Through a thorough analysis of our predictor variables, we obtain a list of variables to use during both logistic and multiple regression. Using the two-step methodology by Sipple (2002), we build thousands of individual regression models and select the 'best' one for each type of regression. The results of these procedures are discussed in Chapter IV.

IV. Results and Discussion

Chapter Overview

This chapter details the results of both our logistic and multiple regression models. We examine how we choose our given models and how we analyze them for statistical validity and user applicability. We begin with the logistic regression analysis before proceeding to the results of our multiple regression analysis.

Logistic Regression Results

As we mention before in Chapter III, building logistic regression models consumes a lot of time. Therefore, we establish a methodology for model building that complements the goal of building a robust model in an efficient manner. Bielecki (2003) calls his approach Darwinist as it replicates the procedure of 'survival of the fittest.' We too use this methodology, though we modify it somewhat. We begin by using JMP® to compute all possible one-variable models. From this initial run, we take forward all one-variable models with a p-value less than 0.2; we now have 15 one-variable models. Using these models, we regress each one against the remaining candidate predictor variables to attain better models. At the end of this round, we select the top 10 models as delineated by the lowest cumulative p-value. We then test the selected models against the remaining predictor variables until the addition of another variable is no longer significant. Finally, to speed the process and reduce the likelihood of developing weaker models, we remove at the end of each stage any variable that does not produce a model with a cumulative p-value of less than 0.20. This step significantly reduces the amount of regressions we run to form our models. However, to safeguard against the exclusion of a

significant variable, we run our final model against all of the predictor variables previously excluded.

In addition to cumulative p-values, we also collect information on the R^2 (U), data point to variable ratio, and the area under the receiver operating characteristic (ROC) curve. While we ensure that our individual p-values remain below 0.05 and the cumulative model p-value below 0.10, the three measures above represent the statistical measures we use to select our model. We summarize these measures in Table 4 below.

Table 4 – Evaluation Measures

Measure
R^2 (U)
Number of Data Points / Ratio
Area Under ROC

We use R^2 (U) as our first level of comparison. As the theory behind logistic regression differs from that of multiple regression, so do the measures. In logistic regression, R^2 (U) represents the proportion of variance explained by the dependent variable, whereas in linear regression it represents the proportion of variance explained by the regression line (Garson, 2003:9). As Bielecki (2003) states quite well, "we consider R^2 (U) as a measure of the amount of certainty explained by our model, and recognize that higher R^2 (U) indicates a better prediction model (55)." For more information on this measure, Sipple (2002) provides further detail.

While we find R^2 (U) to be highly valuable, we need additional measures to find the best overall model. Therefore, we use the data point to variable ratio to ensure that our model, based on a sample of the population, remains representative of the population as a whole. We must watch this ratio because the addition of a variable can do more than

reduce the ratio; the addition of a variable can also remove available data points (as we see below). Neter (1996) states that for every variable present, the model should contain six to ten data points. Based on this recommendation, we follow the precedents of our predecessors and pay close attention to any model that drops below 10:1 and exclude any model that goes below 6:1. By taking this measure, we help ensure that we do not over fit our model to the sample data.

Finally, we evaluate the receiver operating characteristic (ROC) curve as our final measure. Sipple (2002) and Bielecki (2003) examine the ROC curve in technical detail, but for our purpose, we focus on how we implement the information it provides. Specifically, the ROC curve graphs the probability of our model predicting the presence of cost growth when cost growth does indeed exist. As such, the higher the ROC score, the better the likelihood that our model correctly predicts cost growth. Now that we know our evaluation measures, we move to the construction of our logistic model.

Before going on too far though, we note the following quote by Dr. George Box, "All models are wrong, but some are useful." Similarly, model building remains an art as much as a science. While we outline a relatively straightforward process above, which does work for building many thousands of models, our selection of a 'best' model remains subjective and not easily standardized. Indeed, we find that while our process creates valid models to the point of five eight-variable models, we choose as our 'best' model one that began as a hunch while building our five-variable models.

While building our five-variable models, we see that our thirteenth best model (as measured by cumulative p-value) actually produces the highest R^2 (U) by over 0.05 points. While our methodology states that we only carry the top 10 models forward, we

decide to be somewhat subjective (allowing intuition to help guide the process) and carry this model forward under the term ‘Dark Horse.’ Again, at the six-variable point, the Dark Horse remains outside of the top ten, but we believe the additional work to be worth the risk and send it to round seven. In this round, Dark Horse moves into tenth position. After round eight, Dark Horse and only four other models remain; they all meet the p-value cutoffs, have the same data point to variable ratio, and are only separated by ROC curves and R^2 (U). Based on these criteria, we choose Dark Horse as our logistic regression model. As a final step, we also test the Dark Horse model for possible improvement via interaction, higher order terms, and discretizing the continuous variables. The only significant improvement comes as a result of squaring variable #52, *Length of R&D in Funding Yrs*. Tables 5 and 6 below detail the incremental development of the model and Appendix A provides a printout of the model from JMP®.

Table 5 – Logistic Model Performance Measures (Dark Horse)

Measure	Number of Variables								
	1	2	3	4	5	6	7	8	* 8 *
R^2 (U)	0.2164	0.3207	0.3810	0.4461	0.4906	0.5241	0.5551	0.6070	0.6168
ROC	0.8030	0.8615	0.8861	0.9111	0.9263	0.9314	0.9415	0.9541	0.9548
Incremental R^2 (U)	0.2164	0.1043	0.0603	0.0651	0.0445	0.0335	0.0310	0.0519	0.0098
Incremental ROC	0.8030	0.0586	0.0246	0.0250	0.0152	0.0051	0.0100	0.0126	0.0007
Data Points	108.0	108.0	108.0	108.0	108.0	108.0	108.0	105.0	105.0
Data:Variable Ratio	108.0	54.0	36.0	27.0	21.6	18.0	15.4	13.1	13.1
(Note: * 8 * reflects the result of discretizing variable #52)									

Table 6 – Logistic Model P-Values (Dark Horse)

Variable # and Name	Number of Variables								
	1	2	3	4	5	6	7	8	* 8 *
(52) Length of R&D Funding Yrs	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	***
(77) LRIP Planned?		0.0009	0.0011	0.0032	0.0015	0.0015	0.0016	0.0047	0.0070
(64) # Product Variants in SAR			0.0070	0.0062	0.0074	0.0033	0.0038	0.0032	0.0019
(24) Svs>3				0.0081	0.0025	0.0050	0.0026	0.0026	0.0047
(14) Missle					0.0270	0.0128	0.0094	0.0072	0.0113
(62) Proc Started based on Funding Yrs						0.0441	0.0333	0.0224	0.0198
(15) Aircraft							0.0336	0.0131	0.0049
(31) Lead Svc = Navy								0.0371	0.0360
$(52)^2$									0.0003
Cumulative P-Value	0.0000	0.0009	0.0081	0.0175	0.0384	0.0667	0.0843	0.0905	0.0859
(Note: * 8 * reflects the result of discretizing variable #52)									

Using these two tables we determine that all of the measurement increases remain positive and significant at least until squaring variable #52. At this point, the increase appears minor, but we believe it beneficial when combined with the decreased

cumulative p-value. Aside from using a squared term, which adds relatively little complexity to the model, we find no negative aspect of the term's inclusion.

To validate our model, we use the 27 randomly selected programs that we extracted before building our model. These 27 data points represent 20 percent of the original 135 point data set. Using the entire data set, JMP® predicts either a '0' or '1' (no cost growth, yes cost growth) from our model for the remaining 27 data points. Of note, JMP® predicts a '1' for any prediction with a value of 0.5 or greater, while those predictions less than 0.5 receive a '0' (Sipple, 2002). Using this process, we find that JMP® can only predict 25 of our 27 remaining points due to missing data. Nevertheless, our model correctly predicts 19 of the 25 remaining data points for a success rate of 76 percent. While this is not as high as hoped for based off of our performance measures, we assume the model has predictive capability and, due to only losing two data points, has broad applicability to the EMD stage. Table 7 summarizes our validation and Appendix C lists the validation of individual programs.

Table 7 – Logistic Model Validation Results

Validation				
Available	25	of	27	92.6%
Predicted	19	of	25	76.0%

Multiple Regression Results

Now that we possess a logistic regression model, we continue on with Sipple's (2002) two-step methodology by constructing our multiple regression model. Unlike the logistic model which predicts whether or not a program incurs cost growth, the multiple regression model attempts to predict how much cost growth there will be in a program we

envisage to experience cost growth. As we are only concerned with programs that encounter cost growth, our data points reduce from 109 to 78 for model building. For this effort, we again use the predictor variables we use for logistic regression, but we do change our response variable to *EMD Cost Growth %*, which serves as a measurement of the percent increase of procurement cost growth from the Development Estimate.

Since our response variable changes, we perform an analysis on it to ensure that it maintains a continuous nature. Similar to our three predecessors, our response variable exhibits a lognormal distribution and fails to have constant variance via an analysis of the residuals, as well as the Breusch-Pagan test (see Figure 3 below). As a result, we follow the footsteps of our predecessors and transform the response variable using the natural log function. A visual inspection of the distribution indicates it appears reasonably normal, which ensures that our residuals meet the assumption of constant variance by passing the Breusch-Pagan test (see Figure 4 below).

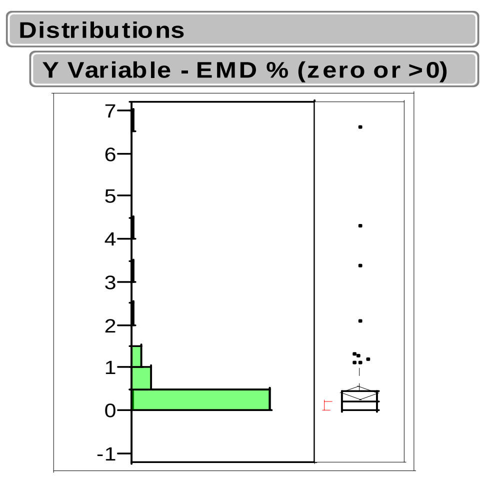


Figure 3 – Preliminary Data Analysis

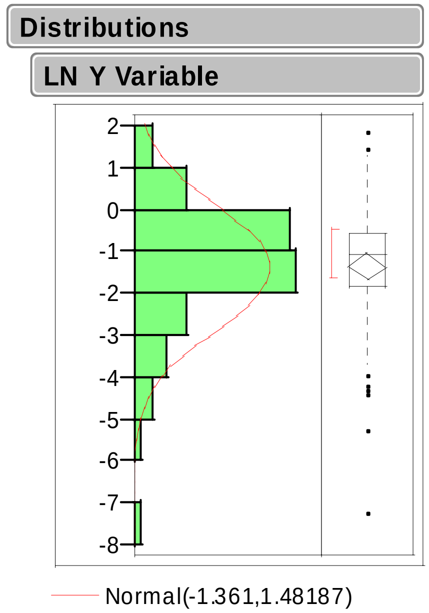


Figure 4 – Transformed Response Variable Results

Similar to the logistic regression process, we build our multiple regression model using the Darwinist approach. We first regress all 77 candidate variables individually and establish our top ten one-variable models. We use these top ten and regress against all remaining variables. We follow this process until our cumulative model p-value no longer remains below 0.10 or our data point to variable ratio drops below 6:1. Using these criteria, we build up to a six-variable model. However, the six-variable model lacks one of our main selection criteria – a high Adjusted R^2 (AR^2), which we use for multiple regression to measure performance rather than the ROC curve and R^2 (U) we use in logistic regression. Indeed, we find that some of our five-variable models contain significantly stronger AR^2, while staying below the cumulative model p-value cutoff and meeting the data point to variable ratio. However, we realize that our most predictive

models contain variable 60, *LRIP Qty Planned*, and the use of this variable reduces our available data points to between 37 and 40 (depending on the given model). As a result, we possess models that may have predictive capability, but also have data point to variable ratios of 8:1 or less. As this could indicate model over fit, we use Cook's Distance on the top couple of models and find that we have multiple data points over 0.50. As a result, we remove them from the model and further reduce our data point to variable ratio. At this point, we believe that our ratio needs to remain much higher to ensure we build a robust model and decide to review our four-variable models for predictive capability. In fact, we find that most of the five-variable models with higher AR^2 originate from one four-model in particular. We show the individual variables of this model and their contributions below in Table 8.

Table 8 – Base Multiple Regression Model

Variable # and Name	Number of Variables			
	1	2	3	4
(48) Funding YR Total Program Length	0.0027	0.0049	0.0043	0.0002
(60) LRIP Qty Planned		0.0046	0.0006	0.0034
(59) Mat of EMD at IOC			0.0414	0.0280
(5) Qty Currently Estimated for R&D				0.0013
Cumulative P-Value	0.0027	0.0095	0.0463	0.0329

We use this model as a point for further testing due to the high AR^2, relatively low cumulative p-value total, and improved 10:1 variable ratio. We test this model as well for positive benefit via interactions and higher order terms, but find nothing significant. We also discretize our continuous variables in hopes of improving our model and find that for variable 59, *Mat of EMD at IOC*, we improve our model by setting a discrete cutoff point. We use the cutoff point to change this continuous variable into a

binary variable that we code as ‘1’ if above a certain value or ‘0’ when below the value. We determine the cutoff value to be 0.90 for variable 59 by analyzing the histogram of the variable and then making minute adjustments to possibly realize further improvement (see Figure 5 below).

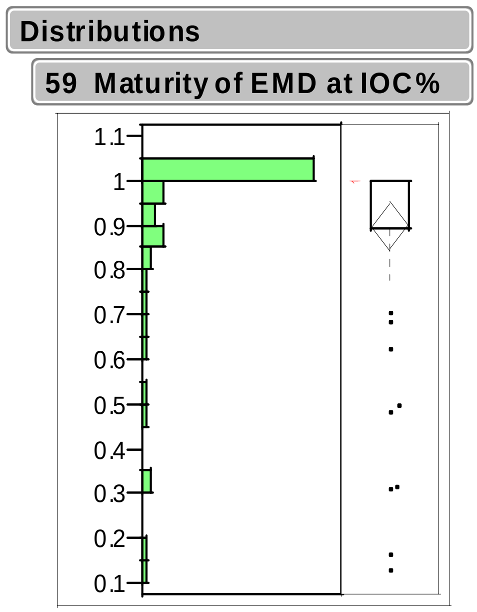


Figure 5 - Variable 59 Histogram

Using this new discretized variable, we select our final multiple regression model. As a final test, we test the model variables for multicollinearity (linear redundancy amongst independent variables) and find that our Variance Inflation Factor scores (all less than 1.2) remain far below the maximum of 10. Tables 9 and 10 below summarize the model and show the development of the model. Appendix B provides the JMP® printout for the model.

Table 9 – Multiple Model Performance Measures

Measure	Number of Variables				
	1	2	3	4	* 4 *
Adj-R^2 (U)	0.1008	0.2176	0.3283	0.4916	0.5228
Incremental Adj-R^2 (U)	0.1008	0.1168	0.1107	0.1633	0.0313
Data Points	78.0	50.0	41.0	40.0	40.0
Data:Variable Ratio	78.0	25.0	13.7	10.0	10.0
(Note: * 4 * reflects the result of discretizing variable #59)					

Table 10 – Multiple Model P-Values

Variable # and Name	Number of Variables				
	1	2	3	4	* 4 *
(48) Funding YR Total Program Length	0.0027	0.0049	0.0043	0.0002	0.0000
(60) LRIP Qty Planned		0.0046	0.0006	0.0034	0.0040
(59) Mat of EMD at IOC			0.0414	0.0280	NA
(5) Qty Currently Estimated for R&D				0.0013	0.0003
(59) Discrete					0.0081
Cumulative P-Value	0.0027	0.0095	0.0463	0.0329	0.0124
(Note: * 4 * reflects the result of discretizing variable #59)					

For validation we use the same 20 percent of the data that we use for logistic regression. While the set contains 27 data points, we find only 6 contain all of the necessary data for our model to use; this leaves only 6 data points for us to use during validation. Using the process of our predecessors, we construct an 80 percent upper prediction bound as we concern ourselves only with positive cost growth. By using an 80 percent upper prediction bound, we validate 5 of the 6 remaining data points correctly for an accuracy rate of 83 percent. We provide Table 11 to summarize our validation and Appendix D lists the validation of individual programs.

Table 11 – Validation for Multiple Regression Model

Validation				
Available	6	of	16	37.5%
Predicted	5	of	6	83.3%

Chapter Summary

In this chapter we build both logistic and multiple regression models in hopes of finding the predictors of cost growth. We select one model for each category that we validate using the data we removed for later validation. Both models contain predictive capability and we believe them to be relatively accurate predictors of cost growth and the total amount of cost growth. We discuss the importance of these findings in Chapter V.

V. Discussion and Conclusions

Chapter Overview

In this chapter, we first compare the results of our study with those of previous and current studies. Then, we summarize our problem statement, study limitations, literature review, and methodology. Finally, we restate our results and provide recommendations concerning future cost growth studies.

Comparison of Predecessor Results

While we primarily focus this research on constructing models that predict the presence and amount of cost growth, other facets remain important. We believe a comparison of our models to those of our predecessors' models one such facet. Specifically, we compare both the logistic and multiple regression models built by Sipple (2002), Bielecki (2003), Moore (2003), and Genest (2004) with our own models in order to determine whether micro or macro models best predict cost growth. Each of the four researchers focuses on different elements within the EMD stage, whereas this study views EMD as a whole. Figure 6 below graphically depicts these differing approaches.

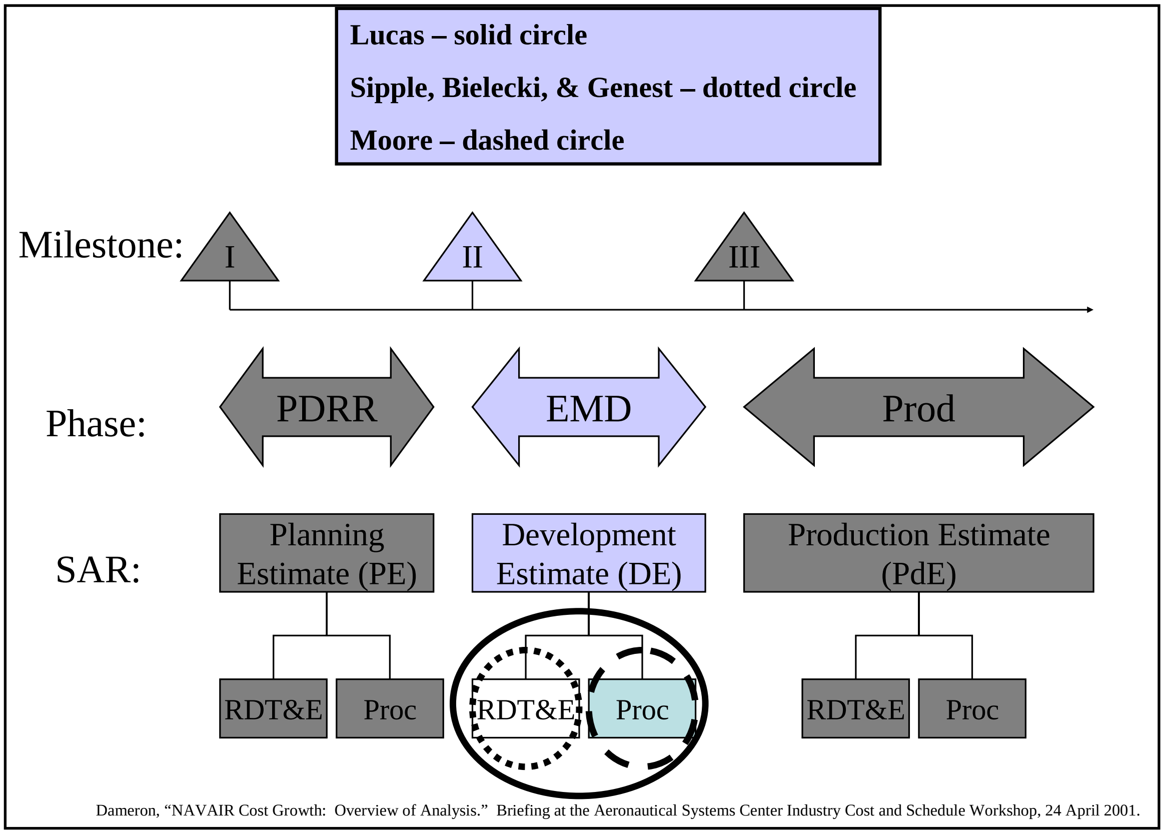


Figure 6 – Prior Approaches to Cost Growth within EMD

While a comparison between these approaches contains limitations, we find the following benefits. First, all five studies use the same database to construct and validate the models, varying only in the total years the database contains (Sipple – 10, Bielecki & Moore – 11, Genest & this study – 12) due to the addition of current data. Secondly, all five studies use the basic methodology Sipple (2002) uses to construct logistic and multiple regression models. Finally, all five studies validate the models with the same process and parameters. Again, the type of funding remains the only change in scope. Tables 12 and 13 document the differences in the ten selected models.

Table 12 – Logistic Regression Model Comparison

Study	Performance Measures				
	R^2 (U)	ROC	Ratio	% Data Available	% Validated
Sipple	0.60120	0.94810	8.70	52%	69%
Bielecki (Est)	0.41840	0.89813	12.60	92%	78%
Bielecki (Sch 2)	0.48080	0.92000	8.75	28%	80%
Moore	0.83070	0.99301	11.60	16%	100%
Genest	0.53570	0.93435	13.70	78%	71%
Lucas	0.61680	0.95478	13.13	93%	76%

Table 13 – Multiple Regression Model Comparison

Study	Performance Measures			
	Adj - R^2	Ratio	% Data Available	% Validated
Sipple	0.42221	14.0	93%	69%
Bielecki (Est)	0.52250	8.8	87%	100%
Bielecki (Sch 2)	0.61900	9.0	91%	80%
Moore	0.52267	8.5	24%	100%
Genest	0.36205	10.0	69%	91%
Lucas	0.52283	10.0	38%	83%

Based on this information, we find no model to be a clear winner. Of the logistic regression models, Moore's model attains the highest performance measure scores and validation percentage, but suffers from the problem of missing data points. One cannot help but question what would happen to the model if those missing data points were available. The next two highest models as rated by their performance measures, Sipple and the logistic model from this study, rank in the bottom half of validation percentage, which further complicates finding a superior model. For cost estimators searching for more macro models, we suggest using the procurement model by Moore if the data can support the model. If not, the overall EMD model from this study may be used in

conjunction with the RDT&E model by Genest; the results should help cost estimators project where cost growth could arise.

Amongst the multiple regression models, Moore's model again validates well, but still suffers from the lack of needed data points. Bielecki's (Est) validates well and attains relatively high performance measures, but the limited scope of his model does not help to predict cost growth on a larger scale. In sum, almost every model built suffers either from a lack of data points, a limited scope, or lower validation percentage. Nevertheless, both of the models by Moore and Genest represent good alternatives as they validate very well. Moreover, the results from these two models may give the estimator more insight than by using the overall EMD model we build as part of this study. However, since all validate above 50 percent, each model may represent a beneficial tool for a cost estimator who can use the model for the given stage of his or her program.

Background of the Problem

Despite numerous reform efforts and constant oversight, cost growth remains a pervasive issue throughout the DoD. When combined with declining budgets, DoD faces the difficult task of selecting between multiple programs that may all be needed. In light of these fiscal woes, cost estimators attempt to predict whether cost growth may occur and how much that growth could be. Most often, cost estimators rely on either expert opinion or historical data to form these estimates, but we provide with this study another lane in an ever-widening avenue. We utilize statistical methods to build predictive

models for cost estimators and thereby provide an objective tool for cost growth measurement.

Limitations

Like most research efforts, our study limits its applicability through the techniques and methodology we choose to use. Specifically, and perhaps the most important limitation, we use a broad range of programs from the SAR and therefore build models that reflect DoD as a whole. Moreover, we differ from prior related research in that we take a more macro view of the database and account for the entire EMD stage of acquisition by combining both RDT&E and procurement funding. As a result, our models may produce errors when used with a program comprised of only one type of funding. However, when used within these parameters, we believe our models predict cost growth with statistical reliability.

Literature Review

To ensure that our study captures the latest developments within the cost field, we perform a literature review of relevant sources. We find Sipple (2002) the most informative and beneficial, but the contributions of Bielecki (2003) and Moore (2003) provide even more explanation about the issue. As such, we focus our efforts on these three studies and pattern our research after them, paying special attention to their methodologies.

Methodology

Our methodology directly springs from that of Sipple (2002). Specifically, Sipple develops the first known two-step methodology to predict cost growth within DoD. Based on his research, we develop both a logistic regression model and a multiple

regression model. The former model predicts 'if' a program will have cost growth and if so, the latter estimates the amount of cost growth to occur.

To build these models, we first update Sipple's original database (which already included the year 2001 data from Bielecki and Moore) to include all relevant data from the 2002 SARs. We only include the latest SAR for each program to ensure independent data points and also convert the data to base year 2002 dollars to account for inflation. In the end, we create a database containing 135 programs, 27 (20 percent) of which we randomly set aside for model validation.

During our preliminary analysis, we discover the same mixed distribution as our predecessors, which further validates our usage of the two-step process. By dividing this mixed distribution through the two-step methodology, we reduce the noise that commonly interferes with the construction of multiple regression models. However, continued data analysis reveals more problems.

Specifically, before we build our multiple regression model, we notice that our response variable forms a lognormal distribution. We follow the procedures Sipple develops and transform the response variable using the natural log function. By transforming the distribution, we achieve an approximately normal distribution with constant variance and construct our multiple regression model.

Results

Our efforts result in two models, one logistic and one multiple. We find our logistic model able to predict 'if' a program will have cost growth 76 percent of the time utilizing 25 of the 27 programs. We find our multiple model to be more predictive, accurately predicting the amount of cost growth 83 percent of the time, but with using

only 6 of the 16 programs due to the missing data variable 59, *Mat of EMD at IOC*, requires. Therefore, we believe the first model to be more universal, but less predictive than the second model. However, we believe that more data for variables 59 and 60, *Mat of EMD at IOC* and *LRIP Qty Planned*, could possibly improve the applicability of the multiple regression model.

Recommendations

The research stream for this stage of acquisition possesses very little water in the reservoir, but reform is like rain – it occurs quite often. Indeed, the milestone category change creates another opportunity for cost growth research. Already, interested parties question how the milestone change from I, II, and III to A, B, and C affects our current predictive models and whether we need new models. In fact, while researching the 2003 SAR database we find multiple examples of the milestone change. Though we do not use these programs in our database because they remain beyond the scope of this study, we believe that once the SARs contain enough new data based on the modified milestones, researchers will again be able to paddle the cost growth rapids and construct models based on the modified acquisition phases.

Appendix A – Logistic Regression Cost Growth Model

Nominal Logistic Fit for EMD (Overall) Cost Growth?

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	37.557442	8	75.11488	<.0001
Full	23.333651			
Reduced	60.891093			

RSquare (U)	0.6168
Observations (or Sum Wgts)	105

Converged by Gradient

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	82	20.037814	40.07563
Saturated	90	3.295837	Prob>ChiSq
Fitted	8	23.333651	1.0000

Parameter Estimates

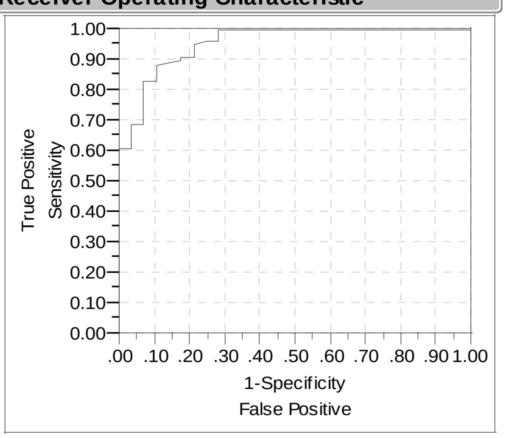
Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	2.00831287	1.406448	2.04	0.1533
77 LRIP Planned?	-3.1573765	1.1707149	7.27	0.0070
64 # Product variants in this SAR	2.28806337	0.7369989	9.64	0.0019
24 Svs>3	6.60259547	2.3329795	8.01	0.0047
14 Missile	3.10471869	1.2255041	6.42	0.0113
62 Proc Started based on Funding Yrs?	-3.6335199	1.5588781	5.43	0.0198
15 Aircraft	7.55927452	2.6840868	7.93	0.0049
31 Lead Svc = Navy	1.94012446	0.9249799	4.40	0.0360
52 SQ (Length of R&D in Funding Yrs)	-0.0224255	0.0062721	12.78	0.0003

For log odds of 0/1

Effect Wald Tests

Source	Nparm	DF	Wald ChiSquare	Prob>ChiSq
77 LRIP Planned?	1	1	7.27361754	0.0070
64 # Product variants in this SAR	1	1	9.63834492	0.0019
24 Svs>3	1	1	8.00953931	0.0047
14 Missile	1	1	6.41822585	0.0113
62 Proc Started based on Funding Yrs?	1	1	5.4328882	0.0198
15 Aircraft	1	1	7.93171522	0.0049
31 Lead Svc = Navy	1	1	4.39941183	0.0360
52 SQ (Length of R&D in Funding Yrs)	1	1	12.7838598	0.0003

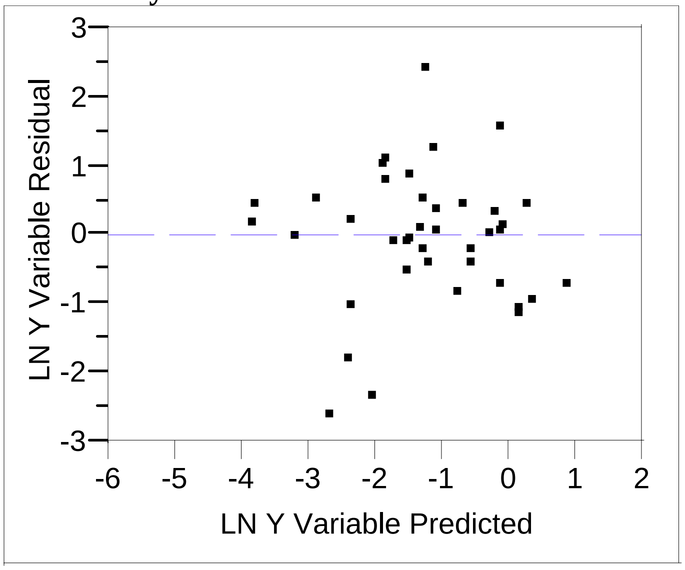
Receiver Operating Characteristic



Using EMD (Overall) Cost Grow th?='1' to be the positive level
Area Under Curve = 0.95478

Appendix B – Multiple Regression Cost Growth Model

Residual by Predicted Plot



Response LN Y Variable

Summary of Fit

RSquare	0.571771
RSquare Adj	0.522831
Root Mean Square Error	1.032355
Mean of Response	-1.27246
Observations (or Sum Wgts)	40

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	49.805021	12.4513	11.6830
Error	35	37.301499	1.0658	Prob > F
C. Total	39	87.106520		<.0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	VIF
Intercept	-2.386329	0.51893	-4.60	<.0001	.
5 Qty currently estimated for R&D	-0.021389	0.005397	-3.96	0.0003	1.1573877
48 Funding YR Total Program Length	0.1062261	0.022202	4.78	<.0001	1.1084036
60 LRIP Qty Planned	-0.000588	0.000191	-3.08	0.0040	1.1260715
#59 Discrete, Mat of EMD at IOC	-1.045019	0.371904	-2.81	0.0081	1.0901402

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
5 Qty currently estimated for R&D	1	1	16.738566	15.7058	0.0003
48 Funding YR Total Program Length	1	1	24.396298	22.8910	<.0001
60 LRIP Qty Planned	1	1	10.133464	9.5082	0.0040
#59 Discrete, Mat of EMD at IOC	1	1	8.414823	7.8956	0.0081

Appendix C – Logistic Regression Model Validation Results

Program	Actual	Predicted	Valid
CMU	1	1	1
AHIP Kiowa Warrior	1	0	0
CG 47 Aegis Cruiser	0	0	1
ATARS	0	0	1
Land Warrior	1	1	1
JDAM	0	1	0
JSIPS TIS	0	0	1
THAAD	0	1	0
Laser Hellfire	1	1	1
Small Missile	0	0	1
RAH-66	1	1	1
BFVS A3 Upgrade	1	1	1
JPATS	1	0	0
AFATDS	0	1	0
DSP	1	1	1
Uh-60M Upgrade	1	1	1
F/A-18E/F	0	0	1
FDS	0	0	1
F-22	1	1	1
MK 50 Torpedo	1	1	1
C-5 RERP	1	0	0
MCS I, II, III	1	NA	NA
E-2C Computer Upgrade	0	0	1
USMC H-1 Upgrades	1	1	1
SBIRS High	1	1	1
FAAD NLOS Fiber Optic Guided-Missile	0	0	1
UH-60A/L Black Hawk	1	NA	NA

Validation		
Available	25 of 27	92.6%
Predicted	19 of 25	76.0%

Appendix D – Multiple Regression Model Validation Results

Program	Actual	Predicted	Valid
CMU	0.1082	-1.094373	1
AHIP Kiowa Warrior	0.3952	NA	NA
CG 47 Aegis Cruiser	0	NA	NA
ATARS	0	NA	NA
Land Warrior	0.06523	NA	NA
JDAM	0	NA	NA
JSIPS TIS	0	NA	NA
THAAD	0	NA	NA
Laser Hellfire	0.68382	NA	NA
Small Missile	0	NA	NA
RAH-66	0.086	0.42754997	1
BFVS A3 Upgrade	0.41026	-2.185556	1
JPATS	0.39207	-0.966771	1
AFATDS	0	NA	NA
DSP	0.21626	NA	NA
Uh-60M Upgrade	0.05758	NA	NA
F/A-18E/F	0	NA	NA
FDS	0	NA	NA
F-22	0.24439	-0.097001	1
MK 50 Torpedo	0.0706	NA	NA
C-5 RERP	0.02658	NA	NA
MCS I, II, III	0.21553	NA	NA
E-2C Computer Upgrade	0	NA	NA
USMC H-1 Upgrades	0.99746	-1.6048351	0
SBIRS High	0.94854	NA	NA
FAAD NLOS Fiber Optic Guided-Missile	0	NA	NA
UH-60A/L Black Hawk	2.42601	NA	NA

Validation				
Available	6	of	16	37.5%
Predicted	5	of	6	83.3%

Bibliography

Bielecki, John V. *Estimating Engineering and Manufacturing Development Cost Risk Using Logistic and Multiple Regression*. MS Thesis, AFIT/GCA/ENC/03-01. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, March 2003 (ADA-413231).

Christensen, David and Carl Templin. "EAC Evaluation Methods: Do They Still Work?," *Acquisition Review Quarterly*, 105 - 116 (Spring 2002).

Coleman, R. L., J. R. Summerville, M. DuBois, and B. Myers. "Risk in Cost Estimating: General Introduction & The BMDO Approach." Briefing at the 33rd Annual DoD Cost Analysis Symposium. Williamsburg VA. 2 February 2000.

Dameron, M. E., C. L. Pullen, J. R. Summerville, R. L. Coleman, and D. M. Snead. "NAVAIR Cost Growth: Overview of Analysis." Briefing at the Aeronautical Systems Center Industry Cost and Schedule Workshop. Wright-Patterson AFB OH. 24 April 2001.

Department of Defense. *Department of Defense Instruction Operation of the Defense Acquisition System*. DoD 5000.2. Washington: GPO, January 2001. Online Defense Acquisition Reference. http://web2.deskbook.osd.mil/reflib/MDOD/012DIDOC.HTM. 3 July 2002.

Drezner, J. A., J. M. Jarvaise, R. W. Hess, P. G. Hough, and D. Norton. *An Analysis of Weapon System Cost Growth*. Santa Monica CA: RAND, 1993 (MR-291-AF).

Garson, David. *PA 765 Statnotes: An Online Textbook.* Textbook for a statistics class taught at North Carolina State University, Raleigh NC. http://www2.chass.ncsu.edu/garson/pa765/logistic.htm. 8 Jan 2003.

Genest, Daniel. *Logistic and Multiple Regression: The Two-Step Approach to Estimating Cost Growth.* MS Thesis, AFIT/GCA/ENC/03-01. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, March 2004 (ADA-TBD).

Gordon, James A. Does A Rubber Baseline Guarantee Overruns? MS thesis, AFIT/GSM/LAS/96S-5. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1996 (AAA-7414).

Hough, Paul G. *Pitfalls in Calculating Cost Growth from Selected Acquisition Reports*. Santa Monica CA: RAND, 1992 (N-3136-AF).

Jarvaise, J. M., J. A. Drezner, D. Norton. *The Defense System Cost Performance Database: Cost Growth Analysis Using Selected Acquisition Reports*. Santa Monica CA: RAND, 1996 (MR-625-OSD).

Jaynes, Howard R, Jr. *Correlation Analysis: Army Acquisition Program Cycle Time and Cost Variation*. MS Thesis. Naval Postgraduate School. Monterey CA. September 1999.

JMP® Version 4.0.4, (Academic), CD-ROM. Computer software. SAS Institute Inc., Cary NC, 2001.

JMP® Version 5.0.1, (Academic), CD-ROM. Computer software. SAS Institute Inc., Cary NC, 2003

Knoche, Chris. "Defense Acquisition Deskbook: Selected Acquisition Report." Online Defense Acquisition Reference Guide. http://www.deskbook.osd.mil/default.asp. 1 July 2002.

Moore, Gary W. *Estimating Procurement Cost Growth Using Logistic and Multiple Regression*. MS Thesis, AFIT/GCA/ENC/03-02. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, March 2003 (ADA-413830).

Neter, John, Michael H. Kutner, Christopher J. Nachtsheim, and William Wasserman. *Applied Linear Statistical Models*. Boston: McGraw-Hill, 1996.

Schick, Allen, and Lastrocco, Felix. *The Federal Budget: Politics, Policy, Process.* Washington D.C.: The Brookings Institution, 2000.

Sipple, Vincent P. *Estimating Engineering Cost Risk Using Logistic and Multiple Regression*. MS thesis, AFIT/GAQ/ENC/02-02. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2002 (ADA-400576).

Weinberger, Sharon. "Programs Must Use Realistic Cost Estimates, DoD Comptroller Says." Aerospace Daily, February 27, 2002. https://www.asc.wpafb.af.mil/news/aero-daily/2002/2002-02-05/ARTICLE198406.html

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	**3. DATES COVERED** *(From – To)*
09-03-2004	**Master's Thesis**	Jun 2003 – Mar 2004

4. TITLE AND SUBTITLE	
CREATING COST GROWTH MODELS FOR THE ENGINEERING AND MANUFACTURING DEVELOPMENT PHASE OF ACQUISITION USING LOGISTIC AND MULTIPLE REGRESSION.	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Lucas, Brandon M., Captain, USAF	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB OH 45433-7765	AFIT/GCA/ENC/04-02
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
ASC/FMCE, AFMC Attn: Mr. Michael J. Seibel 1865 4th St, Rm 134 DSN: 986-5478 WPAFB OH 45433-7123 e-mail: Michael.Seibel@wpafb.af.mil	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Cost growth remains a concern for cost analysts, program managers, senior DoD decision-makers, Congress, and even the American public. All of these people have a vested interest in the cost of DoD programs and most would like to see those costs decrease; as such, we need additional tools to help combat cost growth.

Previous research creates the foundation for the use of a two-step methodology to help predict cost growth, which we follow closely. First, utilizing logistic regression we analyze whether specific program characteristics predict cost growth within the Engineering and Manufacturing Development (EMD) phase for combined RDT&E and procurement budgets. The second step uses this answer (i.e., a positive response) to find cost growth predictor variables. Specifically, we perform a multiple regression analysis and determine the amount of cost growth incurred by these DoD programs. Through these two steps, we seek to unearth any predictive relationships within the data in order to build a predictive cost growth model. The final models predict whether a program will have cost growth and what the potential amount of the cost growth will be for the combined RDT&E and procurement budgets within the EMD phase of acquisition.

15. SUBJECT TERMS
Logistic Regression, Multiple Regression, Cost Variance, Cost Growth, Selected Acquisition Report, SAR, DoD Cost Growth, Cost Growth in DoD Acquisition Programs, Predicting Cost Growth

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
REPORT	ABSTRACT	c. THIS PAGE			Edward D. White, PhD, (ENC)
U	**U**	**U**	**UU**	65	**19b. TELEPHONE NUMBER** *(Include area code)* (937) 785-3636, ext. 4540; e-mail: Edward.White@AFIT.edu

Standard Form 298 (Rev: 8-98)
Prescribed by ANSI Std. Z39-18